

Method, system and medium for detecting and clearing known and unknown computer virus

Publication number: CN1314638

Publication date: 2001-09-26

Inventor: TANG ZHAOMIAO (CN)

Applicant: RUIXING SCIENCE AND TECHNOLOGY (CN)

Classification:

- international: G06F21/22; G06F1/00; G06F9/455; G06F21/00;
G06F21/22; G06F1/00; G06F9/455; G06F21/00; (IPC1-
7): G06F12/14

- European: G06F21/00N3V4T

Application number: CN20011017726 20010429

Priority number(s): CN20011017726 20010429

Also published as:

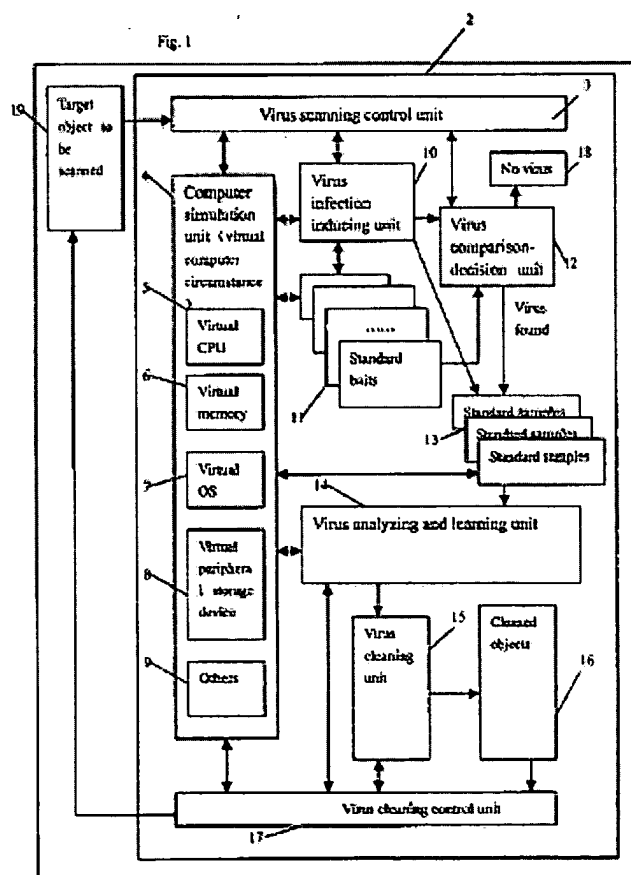


EP1253501 (A2)
US2002162015 (A1)
JP2002342106 (A)
EP1253501 (A3)
CN1147795C (C)

Report a data error here

Abstract of CN1314638

The virus detecting and clearing method includes the steps of: simulation of computer environment; providing several virus infecting objects or lure; loading the detected object to simulated computer environment and activating to induce virus infection and to generate infected standard sample; comparing the infected object with original infected object to judge whether or not there is virus; virus analysis and learning to analyze standard sample and to extract virus related information and knowledge; and clearing virus and correcting the virus modified key information. The present invention can detect and kill known and unknown virus.



Data supplied from the esp@cenet database - Worldwide

Abstract of CN1314638A

Title: Method, system and medium for detecting and clearing known and unknown computer virus

The virus detecting and clearing method includes the steps of: simulation of computer environment; providing several virus infecting objects or lure; loading the detected object to simulated computer environment and activating to induce virus infection and to generate infected standard sample; comparing the infected object with original infected object to judge whether or not there is virus; virus analysis and learning to analyze standard sample and to extract virus related information and knowledge; and clearing virus and correcting the virus modified key information. The present invention can detect and kill known and unknown virus.

1. 一种检测和清除计算机病毒的方法，其特征在于，包括以下步骤：

- 5 计算机模拟步骤，在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境；

提供多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；

装入待检测对象到所述模拟的计算机环境中；

- 10 在所述模拟的计算机环境中激活该待检测对象，以诱发附在所述待检测对象上的病毒对所述多个感染对象进行感染，并生成感染后的标准样本；

将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析，判断有无改变，如有改变，判断所述待检测对象带有病毒；如

- 15 无改变，判断其没有病毒；

病毒分析和学习步骤，用于在所述病毒判断步骤判断有病毒时，对生成的标准样本进行分析，并从标准样本中提取关于病毒的信息和知识；和

- 20 病毒清除步骤，用于根据所述关于病毒的信息和知识，根据病毒对所述感染对象即诱饵的改变，对所述带病毒的待检测对象进行相应的清除病毒体并修正改病毒修改过的关键信息从而清除病毒。

2、如权利要求 1 所述的方法，其中所述计算机模拟步骤包括提供功能函数来调用和执行以下步骤：

中央处理器（CPU）模拟步骤，用于模拟 CPU 的指令；

- 25 操作系统（OS）模拟步骤，用于模拟 OS 提供的各种服务和各种数据结构；

外部存储设备模拟步骤，包括模拟硬盘、软盘等设备的存储空间及结构；和内存模拟步骤，用于生成、分配和管理一模拟的内存空间。

3. 如权利要求 2 所述的方法, 其中所述提供的感染对象包括不同大小不同内容的各种类型的诱饵集, 用于诱发不同类型不同感染条件的病毒, 包括 DOS 文件型诱饵, 用于 DOS com 型文件诱发 DOS com 型病毒; 用模拟的 DOS 引导扇区, 用于诱发 DOS 引导扇区型病毒; WORD 文件型诱
5 饵, 用于诱发宏病毒等等。

4. 如权利要求 3 所述的方法, 其中对同一种类的病毒提供不同大小不同内容的多种诱饵, 尽最大可能的满足待查毒对象里的病毒的感染条件。

5. 如权利要求 4 所述的方法, 其中还包括系统时间模拟步骤, 生成虚拟的系统时间, 用于诱发对时间敏感的病毒。
10

6. 如权利要求 5 所述的方法, 其中所述模拟 OS 包括模拟 DOS, WINDOWS, UNIX 等多种操作系统之一。

7. 如权利要求 1 所述的方法, 在所述病毒清除步骤中, 虚拟执行病毒, 使被感染的宿主对象即被判断为带有病毒的所述待检测对象还原, 从而清除病毒。
15

8. 如权利要求 2 所述的方法, 其中在所述外部存储设备模拟步骤中, 在内存中开辟一块小的内存空间来模拟硬盘, 使内存空间的虚拟硬盘具有跟正常硬盘一样的结构, 包括扇区号、磁道号、柱面号的三维空间概念, 主引导扇区和对应的 0 道中的空闲扇区, 紧接着有引导区, 文件分配表, 根目录区, 还有所需的系统文件, 以及所述用来诱发病毒的诱饵文件等。
20

9. 如权利要求 2 所述的方法, 其中在所述外部存储设备模拟步骤中, 在内存中开辟一块小的内存空间来模拟软盘, 使内存空间的虚拟软盘具有跟正常软盘一样的结构, 包括引导区, 文件分配表, 根目录区, 还有所需的系统文件, 以及所述用来诱发病毒的诱饵文件等。
25

10. 一种检测和清除计算机病毒的计算机系统, 包括一普通的计算机, 其特征在于, 该系统包括:

计算机模拟单元, 在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境;

多个计算机病毒可能感染的感染对象,即诱饵,用于诱发病毒感染;
控制单元,用于装入待检测对象到所述模拟的计算机环境中;

病毒感染诱发单元,用于在所述模拟的计算机环境中激活该待检测对象,以诱发附在所述待测对象上的病毒对所述多个感染对象进行感

5 染,并生成感染后的标准样本;

病毒判断单元,用于将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析,判断有无改变,如有改变,判断所述待检测对象带有病毒;如无改变,判断其没有病毒;病毒分析和学习机,用于在所述病毒判断步骤判断有病毒时,对生成的标准样本进行分析,并从

10 标准样本中提取关于病毒的信息和知识;和

病毒清除单元,用于根据所述关于病毒的信息和知识,根据病毒对所述感染对象即诱饵的改变,对所述带病毒的待检测对象进行相应的清除病毒体和修正改病毒修改过的关键信息从而清除病毒。

11、如权利要求 10 所述的系统,其中所述计算机模拟单元包括:

15 中央处理器(CPU)模拟单元,用于模拟 CPU 的指令;

操作系统(OS)模拟单元,用于模拟 OS 提供的各种服务和各种数据结构;

外部存储设备模拟单元,包括模拟硬盘、软盘等设备的存储空间及结构;和

20 内存模拟单元,用于生成、分配和管理一模拟的内存空间,

所述的各个单元都是一些可调用的功能函数和分配的内存空间,不依赖于具体的 CPU、OS 及外部存储设备。

12. 如权利要求 11 所述的系统,其中所述提供的感染对象包括不同大小不同内容的各种类型的诱饵集,用于诱发不同类型不同感染条件的病毒,包括 DOS 文件型诱饵,用于 DOS com 型文件诱发 DOS com 型病毒;用模拟的 DOS 引导扇区,用于诱发 DOS 引导扇区型病毒;WORD 文件型诱饵,用于诱发宏病毒等等。

13、如权利要求 12 所述的系统,其中对同一种类的病毒提供不同大小不同内容的多种诱饵,尽最大可能的满足待查毒对象里的病毒的感

染条件。

14、如权利要求 13 所述的系统，其中还包括系统时间模拟单元，用于生成虚拟的系统时间，以诱发对时间敏感的病毒。

15、如权利要求 14 所述的系统，其中所述模拟 OS 包括模拟 DOS，
5 WINDOWS，UNIX 等多种操作系统之一。

16、如权利要求 10 所述的系统，所述病毒清除单元虚拟执行病毒，使被感染的宿主对象即被判断为带有病毒的所述待检测对象还原，从而清除病毒。

17、如权利要求 11 所述的系统，其中所述外部存储设备模拟单元
10 在内存中开辟一块小的内存空间来模拟硬盘，使内存空间的虚拟硬盘具有跟正常硬盘一样的结构，包括扇区号、磁道号、柱面号的三维空间概念，主引导扇区和对应的 0 道中的空闲扇区，紧接着有引导区，文件分配表，根目录区，还有所需的系统文件，以及所述用来诱发病毒的诱饵文件等。

15 18、如权利要求 11 所述的系统，其中所述外部存储设备模拟单元在内存中开辟一块小的内存空间来模拟软盘，使内存空间的虚拟软硬盘具有跟正常软盘一样的结构，包括引导区，文件分配表，根目录区，还有所需的系统文件，以及所述用来诱发病毒的诱饵文件等。

19. 一种检测计算机病毒的方法，其特征在于，包括以下步骤：

20 计算机模拟步骤，在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境；

提供多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；

装入待检测对象到所述模拟的计算机环境中；

25 在所述模拟的计算机环境中激活该待检测对象，以诱发附在所述待测对象上的病毒对所述多个感染对象进行感染，并生成感染后的标准样本；

将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析，判断有无改变，如有改变，判断所述待检测对象带有病毒；如

无改变，判断其没有病毒。

20. 一种检测计算机病毒的计算机系统，包括一普通的计算机，其特征在于，该系统包括：

计算机模拟单元，在一台计算机上模拟一个计算机病毒赖以生存的
5 虚拟计算机环境；

多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；

控制单元，用于装入待检测对象到所述模拟的计算机环境中；

病毒感染诱发单元，用于在所述模拟的计算机环境中激活该待检测
对象，以诱发附在所述待测对象上的病毒对所述多个感染对象进行感
10 染，并生成感染后的标准样本；

病毒判断单元，用于将运行后的所述多个感染对象与原始提供的多个感
染对象进行比较分析，判断有无改变，如有改变，判断所述待检测对象
带有病毒；如无改变，判断其没有病毒。

检测和清除已知及未知计算机病毒的方法、系统

5 技术领域

本发明涉及一种检测和清除（即查杀）计算机病毒的技术，特别是一种能检测和杀除未知病毒的方法，以及采用这种方法的计算机系统。

背景技术

10 长期以来，计算机病毒成为困扰使用计算机的人们的一个重大问题。由于病毒的感染性、自我复制性和破坏性，计算机病毒已经威胁到了人们对计算机的正常使用，如造成数据丢失、篡改、文件损坏、软件破坏等。人们经常使用各种各样的杀毒软件来检测和杀除这些病毒。

目前使用的杀病毒软件大都只能检测和杀除已知类型的病毒，也就是，对于已知的各种类型的病毒，已知其特征码，通过对可能带毒的文件进行检测，寻找病毒特征码，如找到相应病毒的特征码，则判断有病毒，进而进行杀毒。这种方法，不能检测到未知类型的病毒。只有在新的病毒被发现并被病毒分析员分析后，才能取得该新病毒的特征码。将这种新的病毒特征码加入到现有的杀毒软件中，才能识别和检测出新的病毒。

自从计算机病毒出现以来，查病毒主要是通过特征值扫描法，即在捕捉到一个病毒后，反病毒人员从病毒的程序体中提取一串或多串特征码，作为该病毒的特征值，杀毒软件根据文件中是否含有病毒的特征值，来判断该文件是否带病毒。虽然十几年来，查毒技术有了一定的改进，但特征值扫描法作为杀毒软件的基础并未动摇。特征值扫描法（即杀毒软件）的最大缺陷是：只有先捕捉到一个病毒并被病毒分析员分析提出病毒特征码加入病毒特征库，然后杀毒软件才能查出该病毒，即杀毒软件总是落后于病毒，病毒一定要让病毒分析员分析。

目前已知的能够检测未知病毒的杀毒技术例如有广谱查毒、启发式

查毒等等，它们是用足够多的经典的病毒特征码，有的还用虚拟机运行待查毒对象的部分代码，运用经验值来判断该待查毒对象可能有毒或判断该待查毒对象里有可疑代码。例如，国内外一些反病毒公司推出了查未知病毒的方法，这些方法都是基于一种思想，总结一些病毒攻击计算机常用的方法，比如写盘，写文件等，然后从查毒对象中查找这些特征，实际这就是行为特征描述。这种方法有的称为诱导式查毒方法，有的称为启发式查毒方法。这种方法可以发现一些未知的病毒，起到警报的作用，但是效果很差，误报率和漏报率都非常高。这有两方面的原因，一是病毒攻击的方法多种多样，很难一一列举，二是病毒采用的攻击方法对于系统来说都是合法的，有很多工具性软件也有同样的行为，很难鉴别。使用这种方法查毒对于用户来说，能查到一些未知病毒，并提醒用户注意，但由于误报率比较高，也给用户带来不必要的恐慌，而且最为重要的一点，它无法杀毒，如果真的被病毒攻击，也只能停机等待杀毒软件的升级。而且这种方法不能确定待查毒对象是否有毒只能说可能有毒。到目前为止，世界上还没有能杀未知病毒的反病毒产品也没有不用病毒特征库（数据库或代码库）能杀已知病毒的反病毒产品。

发明内容

针对上述现有的杀毒软件技术中的问题，本发明的目的在于提供一种能有效检测和清除出已知及未知病毒的方法、系统，它利用病毒的基本特性：感染性，来检测出病毒的存在，从而能从根本上解决检测未知病毒的问题，它能检测出绝大多数已知和未知病毒的存在，进而加以杀除。从根本上改变十几年来每种病毒都必须由人工分析后才能查杀的事实。有了该发明，就使未知病毒能被及时的发现并清除，大大的减轻了病毒对信息及数据破坏的可能性；绝大多数已知和未知病毒的查杀不再需要人工分析，从而节省大量人力和财力。

本发明提供一种检测和清除计算机病毒的方法，包括以下步骤：计算机模拟步骤，在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境；提供多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；装入待检测对象到所述模拟的计算机环境中；在所述模拟

- 的计算机环境中激活该待检测对象，以诱发附在所述待测对象上的病毒对所述多个感染对象进行感染，并生成感染后的标准样本；将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析，判断有无改变，如有改变，判断所述待检测对象带有病毒；如无改变，判断其没有病毒；
- 5 病毒分析和学习步骤，用于在所述病毒判断步骤判断有病毒时，对生成的标准样本进行分析，并从标准样本中提取关于病毒的信息和知识；病毒清除步骤，用于根据所述关于病毒的信息和知识，根据病毒对所述感染对象即诱饵的改变，对所述带病毒的待检测对象进行相应的清除病毒体和修正改病毒修改过的关键信息从而清除病毒。
- 10 本发明还提供一种检测和清除计算机病毒的计算机系统，包括一普通的计算机，该系统包括：计算机模拟单元，在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境；多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；控制单元，用于装入待检测对象到所述模拟的计算机环境中；病毒感染诱发单元，用于在所述模拟的计算机环境中激活该待检测对象，以诱发附在所述待测对象上的病毒对所述多个感染对象进行感染，并生成感染后的标准样本；病毒判断单元，
- 15 用于将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析，判断有无改变，如有改变，判断所述待检测对象带有病毒；如无改变，判断其没有病毒；病毒分析和学习机，用于在所述病毒判断步骤判断有病毒时，对生成的标准样本进行分析，并从标准样本中提取关于病毒的信息和知识；病毒清除单元，用于根据所述关于病毒的信息和知识，根据病毒对所述感染对象即诱饵的改变，对所述带病毒的待检测对象进行相应的清除病毒体和修正改病毒修改过的关键信息从而清除病毒。
- 20 本发明还提供一种检测计算机病毒的方法，包括以下步骤：计算机模拟步骤，在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境；提供多个计算机病毒可能感染的感染对象，即诱饵，用于诱发病毒感染；装入待检测对象到所述模拟的计算机环境中；在所述模拟的计算机环境中激活该待检测对象，以诱发附在所述待测对象上的病毒对所述多个感染对象进行感染，并生成感染后的标准样本；将运行后的所述
- 25 30

多个感染对象与原始提供的多个感染对象进行比较分析,判断有无改变,如有改变,判断所述待检测对象带有病毒;如无改变,判断其没有病毒。

本发明还提供一种检测计算机病毒的计算机系统,包括一普通的计算机,该系统包括:计算机模拟单元,在一台计算机上模拟一个计算机病毒赖以生存的虚拟计算机环境;多个计算机病毒可能感染的感染对象,即诱饵,用于诱发病毒感染;控制单元,用于装入待检测对象到所述模拟的计算机环境中;病毒感染诱发单元,用于在所述模拟的计算机环境中激活该待检测对象,以诱发附在所述待测对象上的病毒对所述多个感染对象进行感染,并生成感染后的标准样本;病毒判断单元,用于将运行后的所述多个感染对象与原始提供的多个感染对象进行比较分析,判断有无改变,如有改变,判断所述待检测对象带有病毒;如无改变,判断其没有病毒。

附图说明

图 1 示出了本发明的检测和清除计算机病毒的计算机系统的结构框图。

图 2A-2C 示出了本发明的检测和清除计算机病毒的方法的工作流程图。

20

具体实施方式计算机病毒之所以被称为病毒,主要是它具有感染性,因此感染性是病毒最本质的特征。如果一个程序具有感染性,就可断定其带有病毒。通过识别一个程序是否具有感染性判断该程序是否被病毒感染,是识别病毒最有效方法。但是,由于病毒具有感染性,要验证它具有感染性,就意味着它要感染某个对象,如果是在真实的环境下进行,这意味着查毒同时,病毒一直在传播,显然不能在真实的环境下进行。验证待查对象是否具有感染性,只能在虚拟环境中进行。

本发明就是利用上述病毒的感染性,把可能带毒的对象放入一个病毒赖以生存繁殖的虚拟的计算机环境里并激活可能带毒的对象,用诱饵

来诱发其感染。并且,由于各种病毒可能会有一定的感染条件,如有的病毒会对目标对象的尺寸、内容等等有一定的要求,因此,本发明提供各种各样的“诱饵”,包括各种各样的不同大小不同内容的诱饵对象,如用 format.com、sort.com 等文件诱发 DOS com 型病毒;用 debug.exe
5 lable.exe 等文件诱发 DOS exe 型病毒;用模拟的软盘引导扇、硬盘一引导扇区、硬盘主引导扇区诱发 DOS boot 型病毒;用 notepad.exe、word.exe 等文件诱发 WINDOWS pe 型病毒;等等。以不同的诱饵对象尽量满足病毒对目标对象的要求。

本发明是指虚拟环境的病毒繁殖—查杀病毒新技术,该技术归属于
10 行为结果查杀毒范畴。本发明用虚拟环境模仿一个真实计算机环境,实现病毒的生存繁殖及传播的全部过程。同时监控病毒的生存繁殖及传播这一过程,并学习病毒的感染方法,进一步推导出病毒感染的逆过程即杀毒方法。具体步骤是如下:第一步,建立一个病毒赖以生存繁殖的虚拟环境,把待查毒的对象也放到该环境里;第二步激活可能带毒的对象,
15 如果待查毒的对象真带毒的话,该虚拟计算机环境也成了带毒环境。对虚拟环境里的诱饵进行各种操作,尽最大努力让病毒感染诱饵。也就是在虚拟环境里做病毒感染繁殖试验,如果诱饵被病毒感染了,说明待查毒的对象真的带毒,被病毒感染的诱饵便成了一个标准样本;第三步,如果上一步感染繁殖试验成功,就用程序来分析标准样本(而不是病毒
20 分析员),并从标准样本中提出查杀该病毒所需的尽可能多的信息。第四步,把程序分析标准样本得到的信息应用到带毒对象上,进一步把病毒给清除。

图 1 具体示出了按照本发明的优选实施例的检测未知病毒的计算机系统的构成方框图。如图 1 所示,在一个普通的计算机系统 1 中,装有一个可由该计算机执行的本发明的查杀毒单元 2。该计算机 1 中可以带有通常的 CPU、内存、操作系统(OS)、外部存储设备(硬盘、软盘等)
25 (图中未示出)。该查杀毒单元 2 的整个程序由计算机 1 的 CPU 中执行。该计算机系统还包括一待查毒对象 19,它可以是计算机 2 的硬盘、软盘中的文件,硬盘、软盘的引导扇区,以及从互联网下载和传输的文件
30 或数据等可能带毒的对象。

如图 1 所示, 所述查杀毒单元 2 包括一查毒控制单元 3, 用于输入待查毒对象 19 到模拟的计算机环境中, 并控制查毒处理各单元的整个过程; 一计算机模拟单元 4, 又称虚拟计算机, 用于生成一个完整的模拟计算机系统, 作为病毒赖以生存繁殖传播的虚拟环境, 该生成的模拟计算机系统可以包括虚拟的 CPU5、虚拟内存 6、虚拟的操作系统 (OS) 7、虚拟的外部存储设备 8 (硬盘、软盘等) 及其它病毒生存繁殖传播所需的系统资源的部分 9, 如系统时间等; 一个或多个标准诱饵 11 (即计算机病毒可能感染的感染对象), 用于诱发病毒感染; 一病毒感染诱发单元 10, 用于将所述待查毒对象 19 装入到所述虚拟计算机 4 中进行运行, 并用所述标准诱饵 11 来诱发待查毒对象 19 可能带有的病毒对这些标准诱饵 11 以及诱发病毒对所述模拟计算机环境本身中的模拟硬盘、模拟软盘的引导扇区等部分的感染, 并生成感染后的标准样本 13; 一病毒比较判断单元 12, 用于检查在诱发病毒前后所述模拟计算机环境本身中的模拟硬盘、模拟软盘的引导扇区等部分有无改变, 以及将所述感染后的标准样本 13 与感染前的标准诱饵 11 进行比较, 检查有无改变, 如有改变, 则判断为该待查毒对象 19 带病毒, 如无改变, 则认为无病毒 18。

所述查杀毒单元 2 的杀毒部分包括一杀毒控制单元 17, 用于控制杀毒处理各单元的整个过程; 一病毒分析学习单元 14, 用于根据所述标准诱饵 11 和感染病毒后生成的标准样本 13 来分析病毒感染所造成的改变, 并学习关于该病毒的知识; 一病毒清除单元 15, 用于根据所述病毒分析学习单元 14 生成的关于病毒的知识, 有针对性地杀除病毒, 并产生清除病毒后的对象 16。该清除病毒后的对象 16 可以由所述杀毒控制单元 17 覆盖到输入的待杀毒对象 19 上, 以消除病毒的存在。

按照本发明的一个实施例, 上述查毒控制单元 2 和杀毒控制单元 17 可以合并为一个控制单元, 以控制上述查毒和杀毒的整个过程。

所述计算机模拟单元 4 生成的虚拟计算机环境包括虚拟机 (虚拟的 CPU) 5、虚拟的操作系统 7、虚拟的计算机外部存储设备 8、虚拟的物理内存 6 等, 总之病毒生存所需的一切计算机资源都被模拟出来; 可能带毒的对象是指理论上说可能被病毒感染的对象。在适当的条件下把可能带毒的对象放到病毒赖以生存的虚拟环境里并在此环境虚拟激活。

所述虚拟 CPU5 又称 `softcpu()` (软件实施或模拟的 CPU)。`softcpu()` 就是一个真实 CPU 指令解释器。`softcpu()` 对程序是解释执行的,就象是一个真正的 CPU,能读懂每一句程序代码并正确解释执行。理论上讲,只要真实 CPU 能执行的代码 `softcpu()` 就能执行,只要真实 CPU 能执行的程序, `softcpu()` 也能解释执行下去;只要真实 CPU 认识的指令, `softcpu()` 也将认识, `softcpu()` 按照真实 CPU 处在这种状态会怎么做,那 `softcpu()` 也照样去做。不过真实 CPU 操作的所有对象(例如: BIOS 芯片、磁盘)是真实的对象,而 `softcpu()` 操作的所有对象(例如: BIOS 芯片、磁盘)都是虚拟的对象。

10 另外, `softcpu()` 只是解释真实 CPU 指令的一个函数而已,可以用汇编语言、C 语言或其它语言写。为了可移植性及可维护性,按照本发明的一个实施例,采用 C 语言编写。

如果要查感染 Intel 计算机上的病毒, `softcpu()` 模拟的是 Intel 的 CPU; 如果要查感染 MAC 计算机上的病毒, `softcpu()` 模拟的是 MAC 的 CPU; 等等。

任何程序总是运行在特定操作系统下,病毒也不例外。虚拟的操作系统 7 就是要模拟病毒所运行的操作系统。该模拟操作系统 7 可以包括病毒运行所需的多个操作系统,如 DOS 的虚拟操作系统、WINDOWS 95 的虚拟操作系统、UNIX 的虚拟操作系统等。为了提高效率,按照本发明的一个实施例,该模拟的操作系统 7 仅模拟病毒所运行的操作系统最小内核。如果要查杀 DOS 病毒,该模拟的操作系统选为 DOS 的虚拟操作系统; 如果要查杀 WINDOWS 95 病毒,该模拟的操作系统选为 WINDOWS 95 的虚拟操作系统,等等。

本发明的计算机模拟单元 4 产生虚拟的计算机外部存储设备 8,包括硬盘、软盘等。在该虚拟计算机环境中,待查毒对象的程序中所有对计算机外部存储设备的读写均是对虚拟外部存储设备的读写,即在程序虚拟运行时,对磁盘文件的感染和破坏数据,都是对虚拟磁盘中的文件感染和破坏虚拟磁盘中的数据。

按照本发明的一个实施例,所述虚拟的计算机外部存储设备 8 包括一可由所述计算机模拟单元 4 调用的虚拟计算机外部存储设备函数或程

序单元 8, 其可生成一虚拟硬盘。该虚拟计算机外部存储设备单元 8 完成的主要功能是在内存中开辟一块所需大小的内存空间, 然后根据具体要求使内存空间的虚拟硬盘具有跟正常硬盘一样的结构, 如有扇区号、磁道号、柱面号的三维空间概念, 主引导扇区和对应的 0 道中的空闲扇区, 紧接着有引导区, 文件分配表, 根目录区, 还有所需的系统文件 (对于 DOS 系统有 IO.SYS MSDOS.SYS COMMAND.COM), 以及用来测试的诱饵文件 (对于 DOS 文件型病毒来说应有例如: DOSEX.E EXE DOSCOM.COM 等文件)。虚拟盘上的对本发明的查杀毒系统有用的数据也不过占几十 K 到几百 K 的字节空间, 而通常的硬盘有几 M 到几 G 的字节空间, 在本系统中其绝大多数空间是没用上的, 所以按照本发明的一个实施例, 可以只在内存中开辟几十 K 到几百 K 的字节空间来虚拟一个几 M 到几 G 的字节空间的硬盘。虚拟一个大的硬盘只需小量内存空间, 所以在普通的微机上就可以实现本系统所需要的虚拟硬盘。并且, 由于在查杀毒期间不用访问实际的硬盘, 所访问的虚拟硬盘实际是在内存的一个小区域中, 所以处理速度快, 节省了时间。而且, 由于该虚拟磁盘仅仅是一块内存空间, 既不会对真实的磁盘造成感染和破坏, 也不会破坏内存的物理特性, 所以对用户的系统是无害的。

按照本发明的进一步的实施例, 在用此单元 8 虚拟硬盘的时候, 可以提供一预定的全局结构 hard-disk-struct 来控制生成硬盘的具体规格, 诸如, 虚拟一空盘, 一引导盘, 一张包含系统文件及诱饵文件的硬盘。

所述虚拟计算机外部存储设备单元 8 也可模拟软盘, 完成的主要功能是在内存中开辟一块所需大小的内存空间, 然后根据具体要求使内存空间的虚拟软盘具有跟正常软盘一样的结构, 如有引导区, 文件分配表, 根目录区, 还有所需的系统文件 (对于 DOS 系统有 IO.SYS, MSDOS.SYS, COMMAND.COM), 以及用来测试的诱饵文件 (如 DOSEX.E EXE, DOSCOM.COM 等), 其所需的数据仅仅只占几十 K。按照本发明的一个实施例, 可以设定一全局结构 floppy-disk-struct 来控制生成软盘的具体规格, 诸如, 虚拟一空盘, 一引导盘, 一张包含系统文件及诱饵文件的软盘, 具体为可以根据该全局变量生成

360K, 720K, 1. 2M, 1. 44M 的软盘。

同样, 可以模拟任何一种操作系统的硬盘和软盘。上述的灵活实现可以节省装入的系统时间开销, 在实际调用过程中, 该虚拟计算机外部存储设备单元 8 把所需数据正确的加载到指定的内存空间。

- 5 上述的虚拟 CPU5、虚拟内存 6、虚拟 OS 7 等的程序单元都是用本领域技术人员已知的编程语言来实现的, 其包含模拟 CPU 的各种指令、内存的各种管理和存取操作、OS 的各种数据结构和功能服务的实现代码, 这些都是目前的编程技术所能够实现的。因此, 这里不再赘述。

- 10 激活待查毒对象就是要让附在待查毒对象上的病毒活动起来并表现出病毒的行为。例如: 如果待查毒对象是可执行二进制文件 (DOS exe 文件、DOS com 文件、DOS bat 文件、Windows NE 或 PE 文件), 那么激活就是执行的意思; 如果待查毒对象是 WORD 等带可执行宏的文档文件, 那么激活就是打开该文档并是能让其中的宏执行的方式打开。

- 15 上述标准诱饵还可设置虚拟的系统时间, 包括各种日期和时间, 用于诱发对时间敏感的病毒, 如类似 CIH 病毒 (在 4 月 26 日发作)、和“黑色星期五”病毒等。如图 1 所示, 本发明的查杀毒程序 2 的查毒部分提供一个标准诱饵的集合, 包括多个标准诱饵 11, 或诱饵集。所谓诱饵, 是指可能被病毒感染的已知对象。按照本发明的一个实施例, 如果
20 要查杀 DOS 病毒, 诱饵是 DOS 程序; 如果要查杀 WINDOWS 95 病毒, 诱饵是 WINDOWS 95 程序; 如果要查杀 WORD 病毒, 诱饵是 WORD 文档; 等等。无论待查杀毒对象是一个什么类型的可执行体, 诱饵都是一个与待查杀毒对象同类型的可执行体。不过诱饵是无毒的, 其尺寸、内容、结构及它的行为功能都是已知, 而待查杀毒对象是否带毒是未知的 (在查
25 之前), 如果真带毒的话, 那么其真实尺寸、内容、结构及它的行为功能都是未知的。

- 30 另外, 上述诱饵 11 不是随便的一个可执行体, 而是经过大量的已知病毒实验, 使其能被大量的已知病毒感染所得到的可执行体。诱饵的尺寸、内容都很合病毒“胃口”即可感染性极强。如果诱饵被病毒感染了, 能够从诱饵中提取出病毒信息。总之, 诱饵就是一个很容易被病毒

感染的已知可执行体, 诱饵集就是多个很容易被病毒感染的各种类型已知可执行体的集合。

具体讲, 按照本发明的一个实施例, 上述标准诱饵 11 的设置包括, 例如, DOS com 型诱饵集, 包含构成诱饵集的多个诱饵文件, 其尺寸分布在 1K 至 60K 之间大小不等的多个文件 (1K、2.5K、12K、20K、30K、40K 等); 诱饵集中文件的第一条指令应分别为 `Jmp`、`call`、`mov`、`xor` 等指令; 诱饵集中文件的时间日期和属性也应分别不同, 用于诱发不同类型的对时间或属性敏感的病毒。

上述标准诱饵的设置还可以包括 DOS exe 型诱饵集, 包括构成诱饵集的多个诱饵文件, 其文件头尺寸分别为 0x20、0x200、0x400、0x600、0x800 等几种, 文件尺寸分别为 4K、10K、20K、40K、80K 等几种, 文件最后一页大小分别为 0x00、0x03、0x80、0x87、0x100、0x198 等几种; 其重定位项数分别为 0x00、0x01、0x02、0x04、0x10 等几种但不占满重定位项表; 其 CS、IP 的值也设为各种各样的值; 程序的堆栈空间分别设有堆栈空间在程序体前部、堆栈空间在程序体中部、堆栈空间在程序体后部、堆栈空间在程序体末端 (不属于程序体) 等几种。

上述标准诱饵的设置还可以包括引导型诱饵集, 其包括: MSDOS、PCDOS、DRDOS、WIN9X 等系统的不同版本的引导扇区及主引导扇区的集合。它实际上是在上述计算机模拟单元 4 生成虚拟硬盘、软盘时, 按照上述引导型诱饵集生成包含各种如 MSDOS、PCDOS、DRDOS、WIN9X 等系统的不同版本的引导扇区及主引导扇区的虚拟硬盘和软盘, 用来诱发引导型病毒。

同样, 上述标准诱饵的设置还可以包括宏病毒诱饵集, 包括各种尺寸、各种类型的 WORD 文件, 用于诱发宏病毒感染。

如图 1 所示, 所述病毒感染诱发单元 10 又称病毒样本生成机, 利用上述提供的各种诱饵集, 进行病毒感染诱发处理, 即运行所述待查毒文件及其中可能带有的病毒, 尽最大努力让标准宿主文件即上述各种诱饵感染上病毒的一个功能单元。染毒判断单元 12 判断是否有任何诱饵在病毒样本生成机 10 中被感染上病毒。具体讲, 染毒判断单元 12 将病毒感染诱发单元 10 中将待查毒对象运行后得到的各个诱饵与运行前的

各自诱饵进行比较,检查有无改变。如果有任何诱饵在病毒诱发运行前后发生了改变,则认定该待查毒对象带毒。同时该改变后的诱饵文件即为病毒样本。也就是说,如果病毒样本生成机 10 没有生成样本 13,说明目标文件无毒;若样本生成机生成了样本 13,说明目标文件有毒并且

5 标准宿主文件(诱饵)成了标准样本,其体内有所有的杀毒信息。按照本发明的一个实例,假设在上述虚拟 DOS 环境中已有 DOS 内存病毒驻留上述虚拟内存,病毒样本生成机对 DOS exe、DOS com 诱饵进行执行、打开、读、关闭、查找等等操作,尽最大努力让内存病毒感染上上述诱饵。被病毒改写、即感染后的待查毒对象即为标准样本 13。

10 对于文件型病毒,感染上病毒的诱饵文件本身就成为上述标准样本。对于引导区型病毒,所述样本生成机 10 根据在上述虚拟硬盘或软盘中的被病毒改变后的引导区信息,生成上述标准样本 13。

所述标准样本 13 是指一个被病毒感染的标准诱饵或指一个被病毒感染的标准宿主。标准宿主是病毒分析员对其大小、内容及体系结构都很清楚的可执行体,其上面很适合附带病毒,如感染条件合适的话。

15

如图 1 所示,按照本发明的一个实施例,本发明的杀毒部分的病毒学习机 14,又称标准样本分析机,它把上述标准诱饵 11 与生成的标准样本 13 进行比较,分析样本,并从标准样本中提取出全部的病毒信息或杀毒所必需的信息。此过程叫病毒学习机的学习过程。病毒学习机的

20 学习过程是一种仿人工杀毒的过程,其不用特征码,与特征码杀毒有本质区别。病毒学习机从标准样本中提取的信息或学习的知识包括:病毒的尺寸大小、病毒在文件宿主中的位置、病毒是否加密变型、病毒是否对宿主进行了加密运算、或病毒已经破坏了宿主程序没法清除(可以删除)、病毒是否对宿主进行了重定位、病毒是否把宿主进行了节对齐、

25 宿主对象的关键信息(例如宿主程序的入口)的值或位置等等。

例如:对于普通 DOS com 型病毒,该病毒学习机 14 提取出两点知识:1.病毒的尺寸大小;2.宿主对象原有功能是否完整或病毒是否已经破坏了宿主程序的原有功能。其中,采用的得到病毒的尺寸大小的算法之一是用标准样本的尺寸大小减去标准诱饵(或标准宿主)的尺寸大小;

30 采用的判断宿主对象原始功能是否完整的算法是在上述计算机模拟单

元 4 中虚拟计算机环境里运行标准样本直到结束或虚拟计算机当机,在此过程中如果产生了标准诱饵(或标准宿主)的原始功能,那么宿主对象原始功能完整,否则宿主对象原始功能不完整。

传统的特征码杀毒方法是用事先由病毒分析员填好的已知病毒特征库里的信息(数据或代码)来清除病毒的。本发明的病毒清除单元 15 是一种仿人工的杀毒单元,它不用病毒分析员填好的已知病毒特征库来杀毒,而是用病毒学习机 14 实时学习得到的病毒知识来把病毒清除。该模仿人工查杀毒的病毒清除单元 15 根据上述病毒学习机 14 提取的关于病毒的信息和知识,把病毒学习机 14 实时学习得到的病毒知识运用到待杀毒对象上并清除病毒。该清毒单元的原则是“解铃还需系铃人”,病毒样本生产机 10 和病毒学习机 14 向病毒学习病毒的感染过程并分析感染的结果(标准样本)得到病毒的数据及属性;病毒的本质在于感染和传播并有一的隐蔽性,也就是说绝大多数病毒不会破坏宿主的原始功能,清毒单元 15 虚拟执行病毒,病毒就会还原宿主对象,清毒单元把病毒还原的宿主对象存盘便可(对于那些在内存与在磁盘的存在形式不一样的对象,要进行相应转换)。至于病毒什么时候还原宿主对象就用病毒学习机学到的病毒属性来判断。例如,按照本发明的一个实施例,采用病毒自还原法就是清毒单元 15 根据感染过程推理出感染的逆过程即解毒过程的方法之一。如果病毒学习机学到的病毒属性或数据足够多的话,就用这些这种属性或数据计算出原始宿主的关键信息(被病毒改过的信息)从而解毒。用软件实时学习并实时运用的杀毒处理在该清毒单元 15 得到实现,而这是目前所有查杀毒软件产品所没有的。

作为一个示例,本发明的清毒单元 15 清除普通 DOS com 型病毒的过程如下: 1. 如果标准样本的原始功能不完整,那么删除待杀毒文件,否则进行下一步; 2. 把待杀毒的 DOS com 文件装入虚拟的计算机环境执行至到虚拟 CPU 程序段寄存器 Cs 值为程序段前缀段地址并且寄存器 IP 值为 0x0100 时为止; 3. 计算待杀毒文件清毒后应有尺寸大小,干净的 DOS com 文件尺寸大小 等于 待杀毒 DOS com 文件尺寸大小 减去 病毒尺寸大小; 4. 生成干净的目标 DOS com 型文件,即把虚拟内存 CS: IP 开始到 CS: IP+干净的 DOS com 文件尺寸大小 的内容存为文件。

如果上述病毒学习机 14 无法学习病毒知识或病毒清除单元 16 判断病毒宿主的原始功能已经被破坏, 则删除所述待查杀对象。

图 2A、2B、2C 示出了本发明的查杀病毒的方法的一个实施例的处理流程图。该流程的各个步骤分别在上述图 1 中的各个处理单元中执行, 以构成完整的查毒和杀毒处理。如图 2A 所示, 首先, 从硬盘、软盘或互联网输入的数据等中待查杀的目标对象 19 的来源中读取待查杀的目标对象 19 (步骤 S101), 然后判断该目标对象是否为可能带毒的对象 (S102)。可能带毒的对象, 顾名思义就是从理论上说该对象有被病毒感染的可能性, 但不一定带毒。可能带毒对象一定是一个可执行体, 例如: *.exe、*.com、*.bat、.doc、PE 和 NE 文件、磁盘引导扇区及主引导扇区, 等等。一个不可执行体不可能是一个可能带毒对象, 例如: *.txt。

如果在步骤 S102 中, 判断所述目标对象是可能带毒对象, 则进至步骤 S103, 进行查杀毒处理; 如判断该对象不可能带毒, 如为不可执行体, 例如 *.txt, 则判断该目标对象无毒; 如目标对象未知, 则报告目标对象为未知对象。

在步骤 S103 中, 由所述计算机模拟单元 4 生成生成一个模拟的计算机环境, 包括生成虚拟 CPU, 虚拟 OS, 虚拟外部存储设备 (硬盘, 软盘), 虚拟内存、虚拟系统时间等, 以便在其中虚拟运行所述可能带毒对象。在步骤 S104, 提供多个计算机病毒可能感染的感染对象, 即图 1 中的标准诱饵 11, 包括上述的文件型诱饵集和所述虚拟硬盘、软盘等的引导扇区诱饵集等。在步骤 S105, 装入所述待查杀对象 19 到所述生成的虚拟计算机环境中。在步骤 S106, 在所述模拟的计算机环境中激活附在所述待查毒对象上的病毒, 即用上述设置的各种诱饵去诱发可能的病毒对上述虚拟计算机环境及各种诱饵文件的感染。在一方面, 在步骤 S107 中, 判断是否有诱饵被感染, 另一方面, 在步骤 S108 中判断该虚拟的计算机环境是否带毒, 即判断所述虚拟的内存中是否带毒, 所述虚拟的各种硬盘的引导扇区是否带毒, 和所述虚拟的软盘的引导扇区中是否带毒。在步骤 S107 中, 如判断有诱饵被感染, 则进至图 2B 中的步骤 S111; 否则, 报告目标对象无毒。在步骤 S108 中, 如判断该虚拟环境

带毒, 则进至图 2B 中的步骤 S110, 对虚拟计算机环境里的诱饵进行尽可能多的操作, 以尽量诱发带毒对各种诱饵的感染, 然后进到步骤 S107, 再判断是否有诱饵被感染。

- 如图 2B 所示, 步骤 S111 报告所示待查毒对象有毒, 产生标准样本, 5 并分析病毒类型, 如该病毒是 DOS 病毒、宏病毒、引导扇区病毒等等。之后进至步骤 S112, 提示用户是否需要杀毒。如用户不需要杀毒, 则给出目标对象带毒的报告, 然后在步骤 S109, 查毒结束。如用户需要杀毒, 进至步骤 S113。

- 在步骤 S113 中, 提取出在虚拟计算机环境中生成的所有标准样本, 10 然后在步骤 S114 中, 利用所述病毒学习机 14 来分析生成的标准样本, 其中主要是判断标准样本的原始功能 (即标准诱饵的功能) 是否完整。在步骤 S115, 判断标准宿主的原始功能 (感染前的功能) 是否完整, 如判断不完整, 则进至步骤 S116; 如判断为完整的, 则进至图 2C 中的步骤 S120。

- 15 在步骤 S116 中, 确定病毒宿主的原始功能已经被病毒破坏无法清除只有删除。在步骤 S117, 询问用户是否删除该带毒文件, 若是则删除该待查毒文件 (步骤 S118), 若否, 杀毒结束 (步骤 S119)。

- 如图 2C 所示, 在步骤 S120 中, 所述病毒学习机 14 从标准样本中学习所有有关病毒的知识, 尽最大努力学到清毒所需的病毒关键数据或 20 属性至到学够了为止。例如: 一个 DOS com 型病毒学到以下知识序列便够 1. 病毒不加密不变形不变长; 2. 病毒的尺寸 virus-size; 3. 病毒只改了宿主的头三个字节; 4. 病毒存放宿主的头三个字节的位位置 data_offset-in-virus (相对病毒体)。

- 然后在步骤 S121 中, 病毒清除单元 15 利用病毒学习机 14 所学到 25 的病毒关键数据或属性在待杀毒对象 (宿主对象) 中的病毒体内查找或计算原始宿主中被病毒修改过的关键数据或属性。例如: 一个有下列属性的 DOS com 病毒 1. 病毒不加密不变形不变长; 2. 病毒的尺寸 virus-size; 3. 病毒只改了宿主的头三个字节; 4. 病毒存放宿主的头三个字节的位置 data_offset-in-virus (相对病毒体)。清除上序病毒步骤 30 如下: 1. 计算病毒体在文件中的位置 virus_offset-in-file 等于 待

杀毒文件的尺寸 `file-size` 减去 病毒的尺寸 `virus-size`; 2. 计算宿主的头三个字节的位置 `data-offset-in-file`(相对宿主文件) 等于 `virus-offset-in-file + data-offset-in-virus` ; 3. 用 `data-offset-in-file` 处的三字节修复宿主文件的头三字节; 4. 把待杀毒文件从尾部砍掉 `virus-size` 字节。

在步骤 S122 中, 判断病毒修改过的宿主信息的原始值是否计算成功。如否, 则杀毒失败 (步骤 S125); 如是, 则进至步骤 S123。

在步骤 S123 中, 修复待杀毒对象(宿主对象)的被病毒修改过的数据或属性, 例如: 文件尺寸、文件头数据等, 这样, 就清除了病毒。

在步骤 S124, 报告杀毒成功。然后, 进至步骤 S119, 杀毒结束。

本发明的上述查杀毒的方法及其各个查杀毒单元都可以通过用普通的计算机编程语言 (如 C 语言等) 编制相应的软件来实现, 该软件可以装在计算机中运行; 可以包含在软盘中进行销售和运行使用; 也可以通过网络及互联网的形式进行传输和下载, 并加以运行。

在本发明的用软件实现的所述查杀病毒的计算机系统和方法中, 能够利用计算机病毒的根本特性: 感染性, 来查出病毒, 并实时学习和实时运用有关病毒的知识, 这是目前所有查杀毒软件产品所没有的。本发明判定病毒的方法是根据“结果”, 而不是根据行为, 可以称其为行为结果判定技术。当然, 本发明的方法也知道病毒的行为, 而且知道病毒的每一种行为和行为的结果, 根据这些行为和结果, 可以安全地把病毒清除, 但是不对单独的行为 (如写盘) 进行判断, 所以可以节省很多时间, 速度较快。并且, 本发明在实际内存中用一小区域提供了病毒生存繁殖的模拟环境, 使得处理速度足够快, 从而使尽量诱发病毒感染能够实际实现。

采用本发明的查杀病毒的计算机系统和方法, 绝大多数已知和未知病毒不再需要人工分析, 不需病毒特征库就能将其查杀; 能及时发现新出现的病毒; 能查杀的病毒数量不再受限制, 有多少查杀多少; 并且, 采用本发明的反病毒软件不再落后于病毒, 能够可靠地发现和杀除未知的病毒。

虽然本发明已以前述优选实施例说明，然其并非用于限制本发明，任何本领域的普通技术人员，在不脱离本发明的精神和范围的情况下，可作各种的更动与修改。因此本发明的保护范围以后附的权利要求为准。

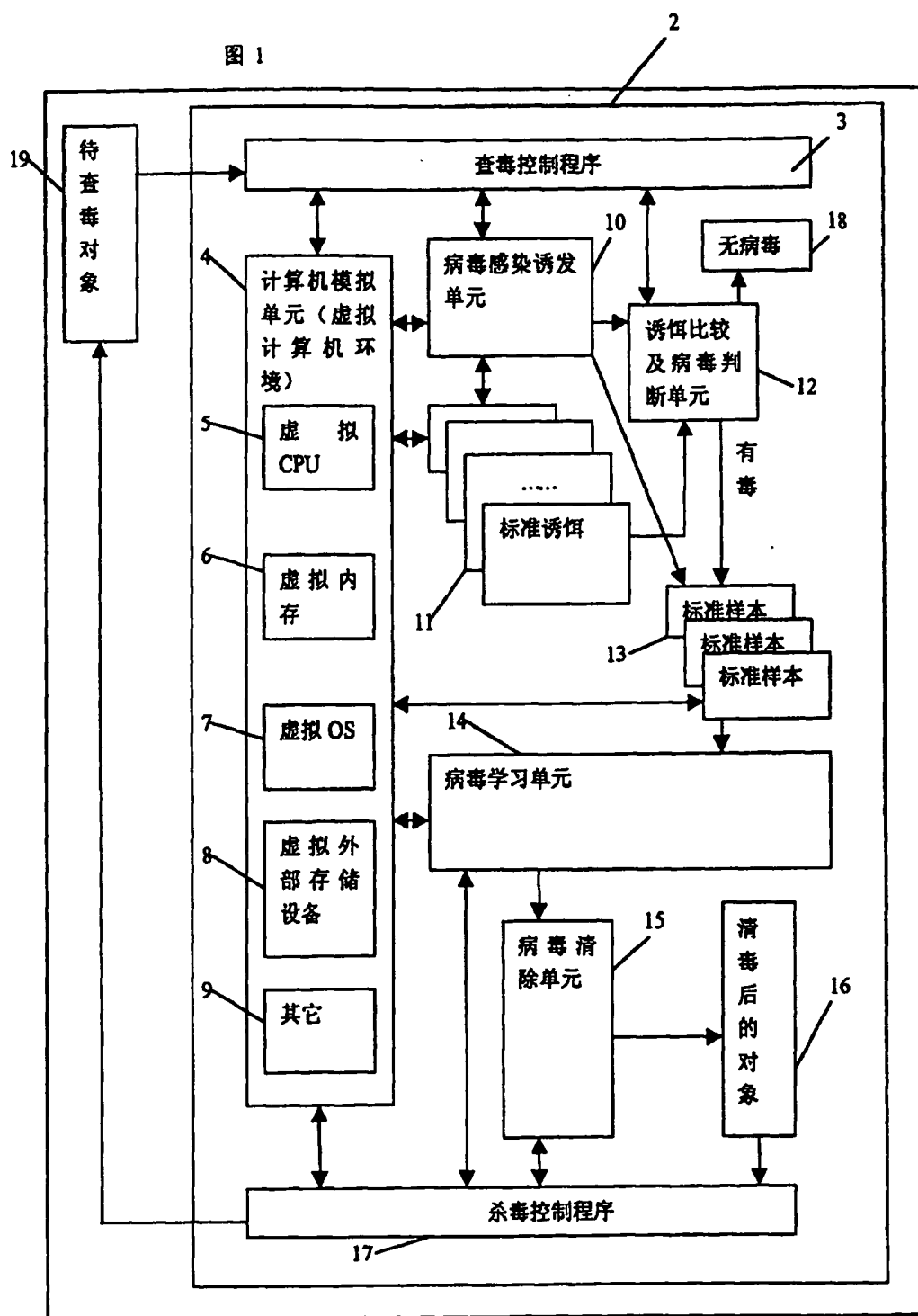


图 2A

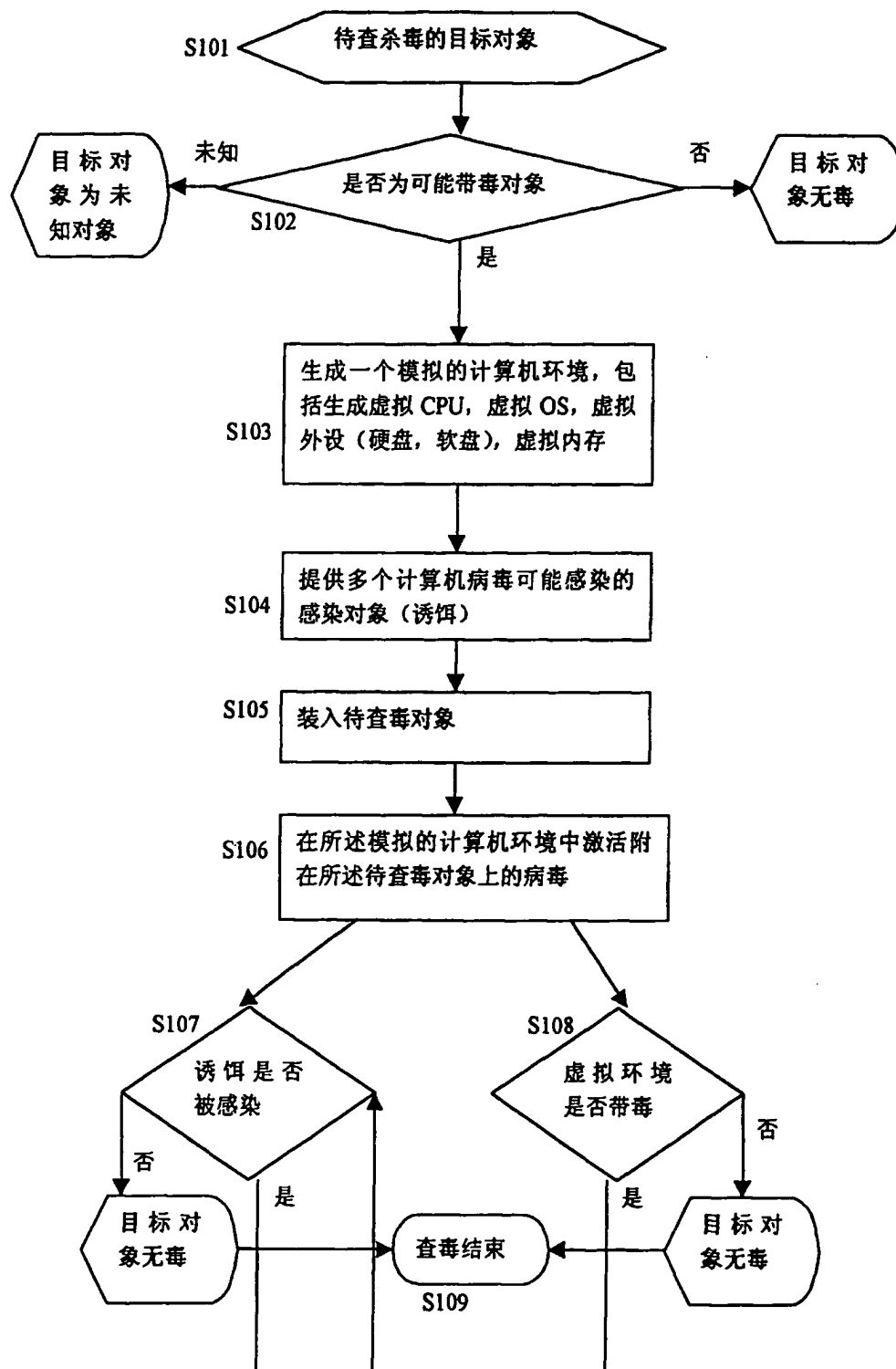


图 2B

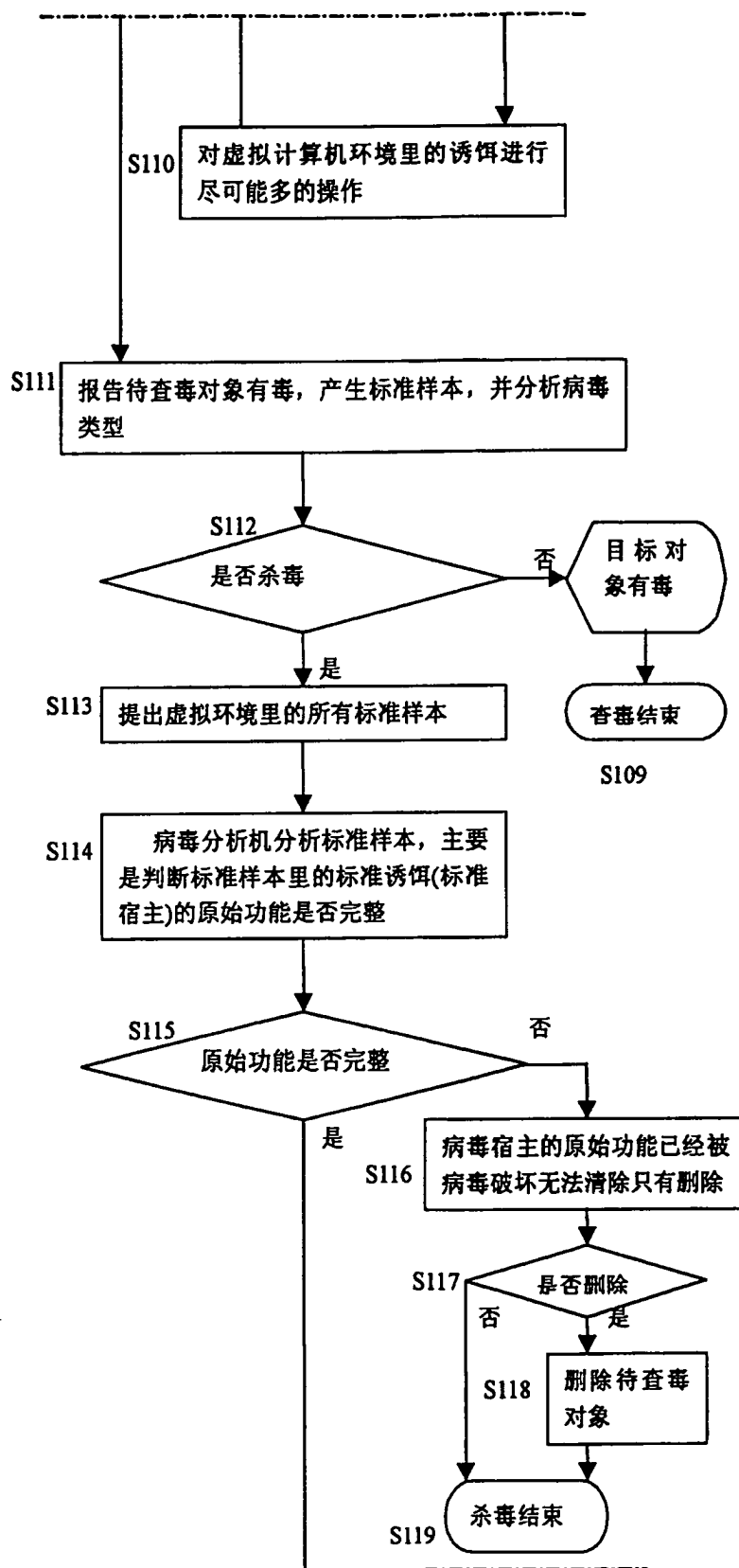


图 2C

